*Kwansei Gakuin University*
*Humanities Review*
Vol. 21, 2016
Nishinomiya, Japan

# Word Association in Japanese Learners of English

Nicholas MUSTY*

Effective English teaching can be well informed by an understanding of the lexical elements of the human mind. Although it is very difficult to uncover the exact ways in which these lexical elements help us communicate, some clues regarding its workings and impact on lexicon can be revealed. One well-known method of accessing the individual lexicon is the word association test (Aitchison, 23-24). This paper reports on such a test, based on McCarthy (152), which was carried out on sixteen learners of English living in Japan.

## I. Stimulus Words

The word association test begins with the selection of stimulus words by a researcher. The following shows the number of appearances in the Bank of English corpus for each of the eight stimulus words, followed by its ranking according to use, in comparison with other words in the test (see table 1). For reference, previous uses of the same words in association tests are also listed, along with the reasons for the selection of each word, including the targeted criteria (McCarthy 152). Section III. 2. explains the process for selecting these words.

**Table 1 List of words selected for the word-association test**

| Word | Frequency | Ranking | Previous use | Reasons |
|---|---|---|---|---|
| Tuesday | 23,859 | 4 | None found | |
| Mouse | 4255 | 8 | Lawson | Homonyms (PC mouse, animal) |
| Rice | 9566 | 6 | Meara | Japanese staple food |
| Exciting | 13,776 | 5 | None found | Emotive but positive |
| Touch | 28,122 | 3 | None found | Well known, many responses possible |
| Back | 344,579 | 1 | Zareva | Possibilities for interpretation are wide |
| Very | 325,336 | 2 | Umemoto | Function word |
| Select | 8558 | 7 | Fitzpatrick | Low frequency, widely known |

* Instructor of English as a Foreign Language, School of Economics, Kwansei Gakuin University

## II. Background to the Test

In investigations of the human mind's lexis bank, the importance of connections has been found, as learners create links between new words and existing ones. Such connections lead to the retention of new input (Sokmen, "Current Trends" 241). One effective representation of the human lexicon which has been suggested is that of a web-like structure of interconnected links (Aitchison, 23-24). This allows for a complex system by which words can be connected to a large number of other words. One theory as to how this works is that a given stimulus triggers a response immediately, which is perceived to be the most relevant response. An alternative to this is that a variety of relevant links are selected initially, with the mind suppressing those deemed to offer little relevance while selecting links considered as most appropriate (Aitchison, 73).

### 1. Word Association Tests

There is evidence of research through this type of test from as far back as the late nineteenth century, when Galton wrote 75 words onto pieces of paper, returning to them only after he had forgotten exactly what he had written, then recording the first words to come to mind from each (Aitchison 23-24). More recent methods generally involve a researcher reading out a list of pre-selected words to a target participant or group of participants, pausing after each to allow for the respondent to give the first word which comes to mind. This answer is then recorded and used to show what kinds of relationship between words are strongest. Although its full value is uncertain, it is thought that this test can reveal some of the inner workings of our lexical processing (Fitzpatrick 320).

### 2. Three General Categories of Association

Patterns of word association have emerged in individual studies as well as across studies. One common distinction between responses has been that of *syntagmatic* responses, in which a participant triggers a word which shares some semantic relations with the stimulus, and generally comes from a different word group. This contrasts with *paradigmatic* responses, which produce an alternative from the same word group. To give an example, the trigger "lion" might generate a syntagmatic response of "frightening" but a paradigmatic one of "tiger". Another kind of answer is known as *clang,* often purely based on phonological aspects. A clang response to lion might be either "lying", which shares phonological features, or "truth", in which case the participant may have misunderstood the input word as "liar" thus uttering what was intended as a (near) antonym. Yoneoka (165) claimed that Japanese and Korean learners of English were comparatively likely to display syntagmatic pat-

terns, although claims of a distinction between L1 and L2 speakers appears to be lacking in evidence, with some studies showing no significant differences in the responses of native speakers and learners (Kruse et al 153).

### 3. More Specific Categories

As well as the three general categories of word association, more specific types of frequent response have also been identified which show more specific connections between words (see table 2).

**Table 2 Syntactic connections between words (Aitchison 87; Carter 20; Fitzpatrick 323)**

| Response type | Example stimulus and response |
|---|---|
| Synonymy | Road/street; Happy/cheerful |
| Antonymy: Complementarity | Male/female; Up/down |
| Antonymy: Converseness | Borrow/lend; Talk/listen |
| Antonymy: Incompatibility | Summer/winter; Orange/blue |
| Antonymy: Graded | Hot/warm/cool/cold |
| Hyponymy: Superordinate/subordinate | Animal/lion |
| Hyponymy: Co-hyponymy | Lion/tiger |
| Collocation | Luxury/hotel; Danger/zone |

These response types obviously overlap in some ways with those of a general nature. For instance, a collocational choice is a kind of syntagmatic choice. There is more emphasis on describing the relationship between words than on trying to explain or guess the reason for the selection, which cannot always be given, even by the respondents themselves (Meara 40).

### 4. Experiential Connections

The associations given so far do not cover all possible responses in the test. Other connections may initially appear to be arbitrary. This comes down to personal experience (McCarthy 41-42; Brown 7-8). Some are cultural, others idiosyncratic or based on knowledge of the source word. In some cases, they might collate strongly (the words “mane” and “lion”, for example, have a t-score of 3.31 in the British National Corpus (BNC) (Walker 103)). Others are less obvious: the response “Prime Minister” to the input of “lion” seems unique at first, but former Japanese leader Koizumi was nicknamed Lionheart due to his hairstyle (McCurry). This kind of co-ordination has been labelled as “encyclopaedic” (McCarthy 41; Brown 7) but I will use “experiential” in order to highlight the relationship between the experience of the participant and the lexical choice expressed.

## 5. Collocations

Collocations, which frequently appear in word association tests (Aitchison 86), may be described as the statistical likelihood of the co-appearance of two words in proximity to one another (Walker 102). A command of collocation is said to be a key indicator of lexical ability (Willis 46). Within the area of collocation, a number of scales can be identified: the amount of restriction in usage, the extent of deviance from standard syntactic structure, and the level of semantic opacity (Carter 70-71). For example, "the lion's share" is quite fixed ("the share of the lion" or "the tiger's share" would not have the same meaning), shows regular syntactic structure, and is semi-transparent in meaning (particularly if the concept of the lion as "king of the jungle" is understood). This paper will explore the kinds of collocational relationships which are demonstrated most frequently, with reference to the BNC, in order to determine whether there is a significant relationship between the two words. This can be indicated by a t-score involving two or more words (Walker 103).

Table 3 shows the most common three collocations for each of the words. I have omitted function words in favour of content words, the latter being able to better demonstrate more unique connections. It should also be noted that the BNC relies heavily on journalism, hence the appearance of political language, obvious here in the collocations for "select".

**Table 3 Common collocations for the stimulus words (from the British National Corpus)**

| Word | Top Three Collocations (content words) in the BNC by t-score |
|---|---|
| Tuesday | night, next, engagements |
| Mouse | Mickey, button, cat |
| Rice | brown, pudding, boiled |
| Exciting | new, prospect, something |
| Touch | get, keep, kept |
| Back | come, go, came |
| Very | good, well, important |
| Select | committee, commons, committees |

Some collocations are obviously cultural and could not be expected to be observed in a Japanese context. For instance, those familiar with the Japanese diet may know that the idea of "rice pudding" would be abhorrent to a lot of Japanese people. Others such as "very good" or "get [in] touch" might be more familiar to English speakers across the globe.

### 6. Word choice

A successful word association test requires careful word selection. Those chosen previously were often idiosyncratic (Meara 40-43) or drawn from the Kent-Rosanoff list (Sokmen, "Word Association Results" 136), which was created in 1910. This includes a lot of quite basic lexis, which is acquired quite early on in the learning process. Therefore, these terms are likely to be more deeply embedded into the participants' language system. It is suggested that words more recently assimilated by learners be used, in order to examine the students' current linguistic system (Fitzpatrick 322). Furthermore, words of high frequency have been found to reveal the least variety of responses ("bread", for example, very often triggers a response of "butter"), while less frequent words ("necessary," "skin") tend to offer more unusual results. Of course, determining what words are difficult or new for a particular learner can also be problematic. However objectively a word list is compiled, inevitably some arbitrary decisions by a researcher are involved.

## III. Method

### 1. Participants

Opportunity was the main factor in participant selection－the sample is one of convenience, consisting of respondents who were easily available to the researcher (Dornyei 129). What they have in common is that they are all adult, study English, and have done so in their free time for at least four years. Their ability ranges from near native to those who could manage only simple conversations, with respondents distributed evenly between these extremes. They all live in Japan and can be considered as fluent Japanese speakers (only one, a Chinese woman who has lived in Japan for most of her adult life, is not native to Japan). There is also one participant, in his 50s, who is recovering from a stroke. Aitchison's (22-23; 88-89) discussion of aphasics suggests that the same problems with lexical recall as non-sufferers tend to prevail in such people, but to a greater extent. However, coordination as a means of connection between words tends to be of greater influence than others.

### 2. Procedure

Despite attempts to create a word list which would not be seen as idiosyncratic, it became obvious that choosing eight words would certainly involve making arbitrary choices. These particular choices were influenced by the work of published researchers, from whom I took some of the words (see table 1). They are not especially academic, as I was concerned that choosing words unknown to beginner students would leave them unable to respond. Each participant was orally given the eight words and asked to reply with the first word that came to mind. As it was an-

ticipated that some words would raise questions as to the exact meaning, I first explained to them that they should interpret the word based on their first instinct.

### 3. Matters Arising from the Test

It can be very difficult to classify responses to a test like this one, as many of the responses failed to fit neatly into any of the categories. For this reason, the experiential category proved to be a rather convenient "catch-all" category. As an example of the usefulness of the experiential category, it can include connections involving the word "back" with "scary", which do not appear to collocate in any significant way, nor do they share a paradigmatic link. In fact, such a test cannot guarantee to fully reveal the mental lexical process, and so it is quite natural that such an apparently opaque relationship would come out of this test. In addition, other connections could be seen to fit into more than one category, and the results show this. To give an example, a hypernym such as "animal" for "mouse" clearly shows both a syntagmatic relationship ("the animal which repulses me most is a mouse") and a paradigmatic relationship ("I saw my cat chasing a mouse/ a small animal") (author's examples). In this respect, the detailed categories were more useful in their concreteness than the three general categories. However, there were other occasions when responses could have fitted into categories other than those given, so inevitably subjective decisions were made. Therefore, making broad generalisations covering all learners is rather problematic, but some patterns will be discussed.

## IV. Test Results and Discussion

The first part of this section will discuss a range of mental links made by participants, including some comparisons with characteristic results as discussed earlier. That will be followed by a brief look at phonology issues (as specified in McCarthy 152).

### 1. General Response Types

Syntagmatic responses featured more than paradigmatic or clang responses, and this difference was somewhat more pronounced at lower levels. However, participant two, who is of an advanced level, only gave syntagmatic responses, making it hard to strongly link this with ability. Otherwise, this pattern tentatively supports the findings of Yoneoka (introduced earlier), suggesting a preference for syntagmatic response. Although research has focused on these categories, their usefulness has not been easy to substantiate here.

**2. Detailed Response Types**

Experiential responses were the most common at the level of detailed analysis (around 35%). This concurs with Meara's (26) declaration that learners exhibit a much wider range of responses than native speakers, often based on quite unexplainable links. The first word, "Tuesday", did not lead to any two participants uttering the same word. Such findings make generalisations impossible and do not offer any obvious insights into the learning process. Co-hyponyms were provided on thirteen occasions, and the stimulus of "rice" produced the most (five occasions). "Curry" was the most commonly given co-hyponym of rice (although this could also be a collocation as "curry-rice" is a popular Japanese dish). Japanese cooking uses rice in a variety of ways, but responses such as "sushi" did not emerge at all. At least from a western perspective, sushi is a highly prototypical Japanese food (with rice). This apparent anomaly might be partly explained by the use of "sumeshi" in Japanese to describe the rice used in sushi, whereas the rice used with curry is known as "raisu", which coordinates with the English word "rice".

**3. Experiential Responses**

Among this group, one pattern which emerged was that of learners at all levels producing adjectives to describe their feelings in response to a noun. These produce such results as "hairy back," "dirty mouse" and "scary (going?) back", apparently a reference to a popular local roller coaster.

**4. Collocations**

As was discussed earlier, many kinds of collocation exist. Correct application of these helps learners to sound more authentic when speaking English, as in the difference between a "strong wind" and a "heavy wind". There was a total of 40 collocations reported here, second only to experiential responses, with the most typical kind of collocation being unrestricted (Carter 70) in nature, typically involving at least one word (either the stimulus or the response) which collocates with many words. Therefore, these do not collocate strongly. One of the least restricted was "touch it", identified here as a collocation due to its appearance in the BNC. Coincidentally, one of the most restricted collocations also emerged from "touch" − "touch and go". Among the sample involved in this study, the collocations which appear to connect most readily are those of relatively free form, while those of a more fixed nature do not emerge so instantly. To aid the learning of related collocations, the introduction of texts containing such patterns as "touch and go", "a select few" or "back to square one", and an exploration in class of these multi-word items would likely have some value. Studying newspaper articles and other authentic texts such as academic papers can help this. The following examples from the BNC demon-

strate how learners might be able to see such patterns in context:

1. "It is still *touch and go* whether Americans will go into the election feeling better off than four years ago or whether they will still be blaming their President for keeping them comparatively poor."
2. "Just five years ago the term desktop publishing was, to all but *a select few,* a totally unknown expression."
3. "If the white-only referendum on March 17 does not back President De Klerk's reforms, then we are *back to square one.*"

After reading a few texts containing such patterns, teachers could write a few collocations on the board and have students work in pairs, trying to create original sentences using them. This would provide a focus on such idiomatic phrases, and would help learners to create stronger connections in their minds between words which have significant links in the utterances of native speakers. Collocations which are frequently produced in native language (see table three) emerged on a few occasions, in particular Mickey Mouse, although this relies somewhat on L1 knowledge. Others to emerge included the concept of "touch" being associated with maintaining friendship, covering "friend" and "see you again," and "very" collocating with adjectives (especially "good"). Further research is needed to compare collocations which emerge from these tests, when they are conducted with both L1 and L2 speakers.

## 5. Reliance on the first language

A reliance on L1 is observable, in particular the stimulus word "mouse," which was responded to by four (25%) participants with "Mickey" and another with "Disneyland". It is not surprising that such responses were more frequent in lower-level students than higher-level ones, as their target language lexicons tend to have fewer connections. However, references to L1 can only be inferred and the data is not strong enough to state this as a certainty.

The findings presented so far tend to support the established literature in this area. There is not enough evidence here to support Fitzpatrick's claim (323) that lower frequency words produce a wider variety of responses – in fact, this paper shows the participants tend to revert to concepts familiar in L1 when a lower-frequency word is chosen. There was also nothing of any significance in the responses of participant seven, who had suffered a stroke. More research could be carried out using a specific class of students and testing them in a similar way on words which had been recently taught to them as new, and comparing these new words with more established ones.

### 6. The importance of phonology at lower levels

Carter's assertion that phonology is of higher importance at lower levels is partially observable here, with several instances of clang. However, they were not limited to the beginner level and can even be seen among advanced students. Due to the nature of the Japanese language, which does not distinguish easily between "b" and "v" or "s" and "th", some of the stimulus words (e.g. "mouse," "very") were subject to misunderstanding. A further study might seek to focus on words such as "vote," "thing" or "lead" in order to establish the extent to which lower-level speakers make phonological associations, as well as the extent to which they simply pick up on the wrong word (as perhaps in the case of "back"), in comparison to more advanced learners.

## IV. Conclusion

A word association test offers a glimpse into the workings of the human mind, but how to interpret its results is not always clear. The results of this test partially confirm the findings of previous research, with evidence that coordination plays a larger role in the word association of lower level learners. An analysis of collocations, a frequent type of response, shows that more fixed patterns are less likely to emerge, and that free collocations are retrieved more readily. Therefore, the learners in this study may benefit from studying texts which allow them to see lexical items in their most natural surroundings and focus on collocations in class. This would obviously help them to sound more natural in their speech. For future study, I would like to suggest the following areas of research:

1. How would word association test results change for these learners if they were given learning opportunities throughout the syllabus which allowed them to focus on some frequent collocations prior to taking the test?
2. What patterns would have emerged if multiple responses were called for as opposed to just one?
3. How big an effect does the selection of words have on results? A longer investigation could test respondents on several words from similar categories, such as words with high frequency homonyms or words which are used as loanwords in the L1, and try to establish patterns more clearly.

It is obvious that the word association test will continue to be explored by linguists for years to come, but it is only when such questions can be answered that language teaching can really start to benefit.

## Works Cited

Aitchison, Jean. *Words in the Mind: An Introduction to the Mental Lexicon* (3rd ed.; Oxford: Blackwell Publishing, 2003).

Brown, Philip Shigeo. *Small-scale Exploration into the Relationship between Word-Association and Learners' Lexical Development* (MA thesis, Birmingham, UK: University of Birmingham, 2006).

Carter, Ronald. *Vocabulary: Applied Linguistic Perspectives* (London: Routledge, 1998).

Coxhead, Averil. "A New Academic Word List." *TESOL Quarterly, 34/2* (2000), pp.213-238.

Dornyei, Zoltan. *Research Methods in Applied Linguistics* (Oxford: Oxford University Press, 2007).

Fitzpatrick, Tess. "Word Association Patterns: Unpacking the Assumptions." *International Journal of Applied Linguistics,* 17/3 (2007), pp.319-331.

Kruse, Heleen et al, "A Multiple Word Association Probe in Second Language Acquisition Research." *Studies in Second Language Acquisition, 9/02* (1987), pp.141-154.

Lawson, Andrew Joseph. *The Relationship between Word-Association and Learners' Lexical Development* (MA thesis, Birmingham, UK: University of Birmingham, 2007).

McCarthy, Michael. *Vocabulary* (Oxford: Oxford University Press, 1990).

McCurry, Justin. "The Lion Weeps Tonight." *The Guardian,* 8 Jul. 1990, http://www.theguardian.com/world/2004/jul/08/worlddispatch.japan . Accessed August 19th, 2013.

Meara, Paul. *Connected Words: Word Associations and Second Language Vocabulary Acquisition* (Amsterdam & Philadelphia: John Benjamins, 2009).

Sökmen, Anita J. "Word Association Results: A Window to the Lexicons of ESL Students." *JALT Journal, 15/2* (1993), pp.135-150.

Sökmen, Anita J. "Current Trends in Teaching Second Language Vocabulary." *Vocabulary: Description, Acquisition and Pedagogy* edited by Norbert Schmitt and Michael McCarthy, Cambridge: Cambridge University Press, 1997, pp.152-161.

Umemoto, Mary. *What We Can Learn From the Word Association of* L2 *Learners* (MA thesis, Birmingham University, UK: University of Birmingham, 2004).

Walker, Crayton. "How a Corpus-based Study of the Factors which Influence Collocation Can Help in the Teaching of Business English." *English for Specific Purposes,* 30 (2011), pp.101-112.

Willis, Dave. *Rules, Patterns and Words: Grammar and Lexis in English Language Teaching* (Cambridge: Cambridge University Press, 2003).

Yoneoka, J. "Word Associations: The Universal Paradigmatic Pattern Assumption and the Japanese/Korean Anomaly," *Kumamoto Gakuen Daigaku Ronshu Sogo Kagaku,* 7-2 (2001), pp.161-174.

Zareva, Alla. "Models of Lexical Knowledge Assessment of Second Language Learners of English at Higher Levels of Language Proficiency." *System,* 33 (2005), pp.547-562.

**Corpora**

The Bank of English (BoE) corpus is jointly owned by HarperCollins Publishers and the University of Birmingham. In 2005, the corpus contained 450 million words. http://www.titania.

bham.ac.uk
The British National Corpus (BNC) is a 100 million word corpus developed in the 1980s. It is maintained and distributed by the Oxford University Computer Service (OUCS). http://www.natcorp.ox.ac.uk

**Appendix**

List of responses by trigger word

| Trigger | Response |
|---|---|
| **Tuesday** | tiring, fire, red, iTunes, work, middle of the week, yes we can, weekday, dancing, fire, Super Tuesday, second, TV programme, grey, Wednesday, melon |
| **mouse** | Mickey, dirty, Mickey Mouse, rabbit, Mickey, computer, peaceful, animal, cat, me, Mickey Mouse, animal, Disneyland, grey, nose, computer |
| **rice** | food, riceball, delicious, ball, Japan, chopstick, chicken, Japanese staple, agriculture, white, curry, food, curry, soft, bread, curry |
| **exciting** | baby, travelling, noisy, amazing, amusement, roller coasters, wonderful, Mount. Fuji, concert, trip, game, enjoy, song, fire, happiness, dancing |
| **back** | front, scary, cannot see, side, big, hair, go, my old days, pain, back monitor, back alright, back yard, travel, flat, forward, good bag |
| **select** | shop, fashion, chocolate, favourite, school, negative, choice, shops, assort, clothes, man, choice, cosmetic, hand, choose, shop |
| **very** | much, beautiful, berry, thank-you, happy, good, nice, expensive, wonderful, strawberry, nice, much, jam, many, extremely, nice |
| **touch** | it, animals, iPod touch, smartphone, and go, hug, see you again, emotional, connection, my dog, out, reach, panel, texture, friend, don't touch me |